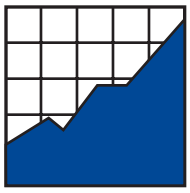
LEP Projects Report 5

Connecting English Language Proficiency, Statewide Assessments, and Classroom Performance



NATIONAL CENTER ON EDUCATIONAL OUTCOMES

In collaboration with:

Council of Chief State School Officers (CCSSO)
National Association of State Directors of Special Education (NASDSE)

LEP Projects Report 5

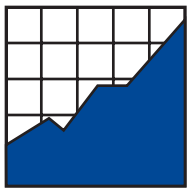
Connecting English Language Proficiency, Statewide Assessments, and Classroom Proficiency

Debra Albus • Jean A. Klein • Kristin Liu • Martha Thurlow

August 2004

All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Albus, D., Klein, J. A., Liu, K., & Thurlow, M., (2004). *Connecting English language proficiency, statewide assessments, and classroom proficiency* (LEP Projects Report 5). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.



N A T I O N A L
C E N T E R O N
E D U C A T I O N A L
O U T C O M E S

The project “Connecting English Language Proficiency, Statewide Assessments, and Classroom Performance” is a subcontract (#41159) with the Minnesota Department of Education supported by a grant (#T 292B010002) from the Office of English Language Acquisition. Opinions expressed herein do not necessarily reflect those of the Minnesota Department of Education or the U.S. Department of Education or Offices within it.

NCEO Core Staff

Deb A. Albus
Ann T. Clapper
Christopher J. Johnstone
Jane L. Krentz
Sheryl Lazarus
Kristi K. Liu
Jane E. Minnema
Ross E. Moen
Michael L. Moore
Rachel F. Quenemoen
Dorene L. Scott
Sandra J. Thompson
Martha L. Thurlow, Director

National Center on Educational Outcomes
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
Phone 612/624-8561 • Fax 612/624-0879
http://nceo.info

The University of Minnesota is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

This document is available in alternative formats upon request.

Executive Summary

In the past, English language learners (ELLs) were expected to learn English before they learned critical content needed to succeed in school. This is no longer the case, especially in light of federal education laws that require that ELLs be held to the same academic standards as other students. At the same time, there are requirements to assess and document progress on tests of English language development. The extent to which these are aligned to tests of academic content and to language proficiency measures is a matter of debate, with little relevant research available.

This study sought to provide information on the links between academic language, language proficiency tests, and performance on standardized assessments by examining relationships among: (1) two language proficiency measures (e.g., Language Assessment Scale (LAS) and Minnesota's Test of Emerging Academic English (TEAE), (2) teacher ratings of classroom reading and writing samples, and (3) two state achievement tests: Minnesota's Comprehensive Assessments (MCAs) and Minnesota's Basic Skills Test (BSTs). The students in this sample were 99 English language learners (ELLs) in grades 3, 5, and 11.

Language Assessment Scale (LAS) and Test of Emerging Academic English (TEAE)

In examining the relationship between the LAS and TEAE for our sample of students, we found that the underlying reading skills being measured by the two tests were closely related. Of course, this does not mean that the tests are measuring the exact same skills. Instead, it indicates that the students who performed one way on one test tended to perform a certain way on the other test. The writing tests for the LAS and TEAE were not related, indicating that the tests are either measuring different skills, or are measuring underlying skills differently.

Teacher Ratings of Reading and Writing

Teachers who were licensed to teach English as a Second Language (ESL) tended to rate students' skills as either the same or higher than the students' content teachers. But both ESL and content teachers tended to rate listening and speaking skills higher than students' skills in reading or writing. Teacher opinions about specific students' chances to succeed in future classes without further language support were related to whether students had achieved passing scores on the Basic Skills Tests and had at least been rated as "achieved" on the Minnesota Comprehensive Assessments.

Proficiency Tests, Teacher Ratings, and State Tests

Although certain Language Assessment Scales scores (reading and overall) were significantly correlated with the MCA writing test and BST reading test, there appeared to be stronger correlations between the TEAE reading score and the MCA and BST reading scores. A modest correlation was found between the TEAE reading and MCA writing test. The TEAE writing score had a strong correlation with BST reading, though the number of students on which this

correlation was based was small (17). ESL reading and writing samples were also correlated with the BST reading test, however these numbers of students on which these were based were also small (10 and 11).

Comparisons of ESL and content teacher ratings of student ability with scores on state achievement tests showed inconsistencies based on whether the teacher was an ESL or content teacher. ESL teacher ratings generally increased with increasing ability in relation to student performance on standardized tests, whereas content teacher ratings did not show this relationship consistently. These inconsistencies may have existed for several reasons, including potential differences in how ESL and content teachers interpreted students' skills using a rubric designed to measure language development, and differences in how classroom activities were able to capture language and content development.

Limitations of Study

The findings of this study are limited to some extent because of the difficulty encountered in recruiting adequate numbers of subjects and teachers willing to participate. Reluctance to participate was encountered despite generous subject and school incentives. Other study limitations included the lack of a complete state database which made it difficult to acquire test scores for students. Difficulty in getting ESL and content area work sample ratings from teachers further complicated completion of the study and suggests that accountability measures that rely on work samples may have similar problems. A small number of students also were quite familiar with the oral component of the LAS, an unexpected situation that could confound results.

Classwork samples, as intended in the study design, were not standardized. Even though staff and teachers rated samples with the same rubric, questions must be raised about the reliability of ratings. It is helpful to examine the classwork sample ratings in the context of how they related to other more objective measures of language proficiency and achievement.

Conclusion

The comparisons of proficiency test data and state test data showed that certain language proficiency scores (reading and overall) were significantly correlated to the MCA writing test and BST reading test. However, there appeared to be stronger correlations between the TEAE reading score and the MCA and BST reading scores. In comparing student performance on state tests with teacher ratings on a question about students' ability to succeed without further language support, most of the students rated as likely to succeed had achieved passing scores on the BST and had at least been rated as "achieved" on the MCA tests. Still, sometimes a student with lower teacher ratings had also achieving passing scores on state measures such as the BST Reading test. Other findings of interest included differences in how ESL and content teachers rated student skills and abilities.

Overview

Historically, much of the instruction of English language learners has been oriented toward moving students to a certain level of English proficiency before academic content is introduced. It has been understood from the research literature (Collier, 1992; Cummins, 1989; Ramírez, Yuen, Ramey, & Billings, 1991) that English Language Learners (ELLs) cannot wait for language skills to fully develop before these students are taught the crucial academic content required of all students to succeed in our school systems. Instead, these students must master a full range of social and academic uses of English at the same time they are leaning reading, mathematics, and other content area knowledge. The *No Child Left Behind Act* of 2001 (NCLB) has given this understanding legal footing by mandating that English language learners be held to the same academic standards used for structuring what is taught and assessed for all students in addition to learning English language skills (i.e., listening, speaking, reading, writing, and comprehension) based on English language proficiency (ELP) standards set by each state.

It follows that these ELP standards and the associated annual English proficiency measure should reflect development in the academic language skills needed to successfully develop in academic content areas. Yet Title III legislation does not explicitly require the measurement of "academic language" in the annual language proficiency measure. The language in Title III non-regulatory guidance (2003) instead requires that a state's ELP assessments be aligned with ELP standards that are linked to the regular content standards. This is evident in several statements in the guidance:

> Although English language proficiency and reading/language arts academic standards are different, they should be linked to one another. English language proficiency standards should define proficiency levels that will help LEP students to acquire the English language skills necessary to meet academic content and achievement standards. (p. 8)
>
> States are encouraged, but not required, to align English language proficiency standards with academic content and achievement standards. (p. 9)
>
> English language proficiency assessments must be aligned with English language proficiency standards and provide a means of demonstrating progress towards meeting the English language proficiency annual measurable achievement objectives. (pp. 9-10)

The Title III guidance highlights the difference between alignment and linking of academic content to ELP standards. The extent to which the ELP standards and assessments are linked to regular content standards and assessments is assumed to affect the progress of English language learners toward academic target skills and abilities through the educational system.

Yet the process of linking and aligning standards of language proficiency to classroom and achievement assessments is not uncomplicated. Instead, the task of defining academic language is quite complex. A few researchers have contributed to the field by fleshing out the definition that academic language is language used in an academic setting in order to acquire knowledge. Stevens, Butler, and Castellon-Wellington (2000) approached the task of defining academic language by incorporating three main perspectives: (1) theories based on language functions and structure (Chamot & O'Malley, 1994), (2) distinctions between socially-based communications referred to as basic interpersonal communicative skills (BICS) and more abstract cognitive academic language proficiency (CALP) (Cummins, 1984), and (3) a sociolinguistic view (Solomon & Rhodes, 1995) that defines academic language contextually according to register (degrees of informal and formalness) for specific tasks.

Scarcella (2003) used a similar "define by combining" approach, noting the BICS/CALP distinction made by Cummins, but also a perspective not directly mentioned by Stevens et al. (2000), which emphasizes the concept of multiple literacies. This second perspective is distinguished by its allowance of academic language to be dynamic, evolving, and accepting of the students' own expression of academic ideas rather than focusing on the exact form in which they are expressed. Scarcella advocates, like Stevens, that academic language involves linguistic and dynamic competencies.

Therefore, even with the most well-crafted English language proficiency standards, the practical task of applying an operational definition of academic language proficiency to the classroom or attempting to measure its incremental development is a challenging one for states. States have temporarily resorted to the use of commercially available assessments originally designed for placement and diagnostic decisions to meet the need of federal legislation requirements, while testing companies work with states to create new standards-based assessments focusing on academic language proficiency (Olson, 2002). Among the most widely used commercial assessment is the Language Assessment Scales (LAS) test developed by DeAvila and Duncan (1990). Although this test was originally intended for placement decisions, its publisher has signaled the probability that future changes may be made to better reflect the needs of states dealing with requirements to link the proficiency measure to state standards (Jackson & Jackson, 2003). Some states have developed their own versions of language proficiency assessments in an attempt to capture academic language development aligned with state ELP standards. Minnesota is a state that has done this with its language proficiency assessment called the Test of Emerging Academic English (TEAE).

In theory, annual assessments of academic English proficiency development should result in more accurate information about the progress of students in language and content knowledge. Yet limited research has been conducted to demonstrate that this is the case as indicated by Stevens et al. (2000) who argued that "little research exists on the relationship between academic

language, language proficiency tests, and performance on standardized assessments" (p. 4). We do know from research studies that language proficiency may account for between 16% to 50% (average 25%) of the variation in content assessment scores (Stevens et al., 2000) depending on the language load of a content test (for example, mathematics versus social studies). We also know that the language of standardized tests may contribute to their difficulty for English language learners (Bailey, cited in Stevens et al., 2000). Further, Cunningham and Moore (1993) and Abedi, Leon and Mirocha (2001) found improved performance for ELLs when the language load of test items was reduced. Still, there remain questions about differences that are assumed to exist between the language used and measured on language proficiency tests and the language used on large-scale content assessments.

The purpose of this study was to examine how two language proficiency tests relate to: (1) classroom performance in content areas, and (2) standards-based content assessments. One language proficiency test was the LAS, originally designed for instructional placement of ELLs. The other was the TEAE, which claims to capture academic language aligned with the English proficiency standards in Minnesota. Our goal was to describe the role that academic language might play in determining differences among language proficiency tests.

Research Questions

- How does student performance on the LAS correlate to their performance on the TEAE?

- How does LAS and TEAE performance compare to educator ratings of classroom work?
 - How did ESL and content teachers rate student skills across modalities?
 - How did ESL and content teachers rate student skills related to academic work?
 - How did teachers rate work samples using the TEAE linked rubric?
 - How did language proficiency assessment performance correlate to teachers' ratings of student work samples?

- How does LAS and TEAE performance compare to statewide achievement test data in reading and writing?
 - How did teachers rating of student work samples compare to the proficiency assessments and state achievement tests?
 - How did students rated likely to succeed do on language proficiency and state achievement tests?
 - How did students' language proficiency assessment performance and work sample ratings correlate to state achievement tests?

- How did students with very low oral scores on the LAS test do on proficiency assessments, work samples, and state achievement tests?

Method

English language learners and their teachers were recruited from three school districts in Minnesota. The students, all of whom were participants in the state TEAE, were tested by the project using the full Language Assessment Scales (LAS). Teachers also provided information about classroom performance by collecting and rating classroom samples.

Data were analyzed in the following order: (1) student performance on the LAS compared to the TEAE, (2) TEAE and LAS performance compared to ESL and content area teacher ratings of collected classroom samples, and (3) TEAE and LAS performance compared to other statewide achievement test data in reading and writing. In this section we describe in more detail the research participants, measures, and procedures.

Participants

The student participants in this study were 99 English language learners in 3rd, 5th, and 11th grade from one urban and two suburban school districts in Minnesota. By grade, there were 35 fifth graders, 31 eighth graders, and 33 eleventh graders who participated.

Over 20 districts were contacted for possible participation between the spring of 2003 and spring of 2004 for potential participation. Aside from the three districts in our study, all other districts (with large enough ELL populations) that were contacted declined to participate. Among these, eight did not respond to phone calls. The study design called for representation outside the metropolitan area. However, of 11 districts outside the metropolitan area that had ELL populations large enough to contact, one agreed but then withdrew due to lack of teacher interest.

Table 1 describes the participants by grade, gender, and language. Table 2 summarizes the overall gender of participants.

We provide descriptive information about the districts and schools participating in this study in Tables 3 and 4. Table 3 gives information on total student population size, percent receiving free and reduced lunch, percent ELL, and percent increase in ELL student population since 2000. Table 4 provides similar information for each participating school. At the school level, the percentage of students receiving free and reduced lunch ranged from 17% in Suburban 1 to 78% in the Urban district. The percentage of ELLs in each school ranged from 6% in Suburban 2 to 36% in the Urban district.

Table 1. Participating Students by Grade and Language Across Sites

Languages	Urban			Suburban 1			Suburban 2		
	Gr. 5 N= 3	Gr. 8 N=5	Gr. 11 N=14	Gr. 5 N=16	Gr. 8 N=17	Gr. 11 N=12	Gr. 5 N=16	Gr. 8 N=9	Gr. 11 N=7
Hmong	33 %	20 %	14 %	69 %	47 %				
Somali	66 %	60 %	57 %			8 %			
Spanish			21 %	25 %	18 %	50 %	75 %	44 %	57 %
Russian					6 %	8 %	13 %	33 %	29 %
Other*		20 %	8 %	6 %	24 %	33 %	19 %	22 %	14 %

* Students include: Amharic, Bamanan, Cambodian, Creolized English, Laotian, Oromo, other African, and Vietnamese.

Table 2. Participating Students by Grade and Gender Across Sites

Gender	Urban			Suburban 1			Suburban 2		
	Gr. 5 N=3	Gr. 8 N=5	Gr. 11 N=14	Gr. 5 N=16	Gr. 8 N=17	Gr. 11 N=12	Gr. 5 N=16	Gr. 8 N=9	Gr. 11 N=7
Female	3	3	6	10	7	7	5	7	4
Male	0	2	8	6	10	5	11	2	3

Table 3. Demographic Information for Participating Districts and the State of Minnesota

For 2003-2004	Urban	Suburban 1	Suburban 2	State
Total Student Population	45000	14000	4500	830000
Percent Free and Reduced Lunch	68 %	30 %	21 %	28 %
Percent ELL	23 %	8 %	12 %	7 %
Percent Increase in ELLs since 2000	21 %	139 %	116 %	52 %

Table 4. Demographic Information for Participating Schools

For 2003-2004	Urban		Suburban 1			Suburban 2		
Grade	K-8	9-12	K-5	6-8	9-12	5-6	7-9	10-12
Percent Free and Reduced Lunch	23%	78%	64%	39%	17%	29%	24%	20%
Percent ELL	10%	36%	31%	13%	8%	12%	8%	6%
Size of School	600	1300	500	1400	2300	700	1100	1000

Teachers in the study were either ESL teachers (N=14) or content area (i.e., science or social studies) teachers (N=17). They participated in this study by collecting and rating student classroom samples for three students per school site. Some teachers rated more than one student. Complete information on the teachers' backgrounds is presented in Tables A1 and A2 in Ap-

pendix A. Table 5 shows summary information for these teachers by type of licensure (e.g., ESL or content), licensure for current grade he or she is teaching, and years teaching. Based on teacher self-report, all teachers were qualified to teach their respective areas and grades. Further, half of the teachers, both ESL and content, had taught at least four or more years. Although not shown in Table 5, most teachers had spent these years teaching in their current district (see Appendix A).

Table 5. ESL and Content Teacher Licensure Information and Years Teaching

	Urban		Suburban 1		Suburban 2	
	ESL N = 4	Content N = 4	ESL N = 6	Content N = 5	ESL N = 4	Content N = 8
Percent Licensed in ESL	100%		100%		100%	
Percent Licensed in Content		100%		100%		100%
Percent Licensed in Appropriate Grade	100%	100%	100%	100%	100%	100%
Years Teaching:						
More than 7	25%	100%	17%	80%	25%	25%
4-7	75%		33%		50%	25%
1-3			33%	20%	25%	25%
Less than 1			17%			25%

Language Proficiency Measures

LAS

We administered the reading and writing Language Assessment Scales (LAS-R/W Forms B) and the oral Language Assessment Scales-Oral (LAS-O Form D). The LAS-R/W, a battery of reading and writing competency tests, and the LAS-O represent a "convergent approach" to assessing language (Duncan & DeAvila, cited in Del Vecchio & Guerro, 1995), which places the most importance on the combined total score for use in identification and instructional decision making. The current LAS test combines discrete point items with holistic measures in both the LAS-R/W and LAS-O. We used the long form of the LAS-O, which included a minimal-pairs section with measures of vocabulary, listening comprehension, and story retelling.

Del Vecchio and Guerro (1995) have noted that more recent descriptions of the LAS include the ability to forecast the likelihood of student success in a mainstream classroom (Del Vecchio et al., 1995). LAS authors have further made known their intentions to make the LAS test more finely tuned to the purpose of assessing language proficiency aligned with state English language development standards (Jackson & Jackson, 2003). These changes are occurring, no doubt, in response to recent federal requirements for the assessment of ELLs, and the relationship this assessment must have directly to ELD standards and indirectly to states' content standards.

TEAE

We also administered the Test of Emerging Academic English (TEAE), which is a Minnesota adaptation of the Illinois Measure of Academic Growth in English. It was developed to measure the growth of the English language learner's academic proficiency in English and had been linked to Minnesota's content standards. It is currently administered to fulfill Title I accountability requirements as an annual measure of growth in English language proficiency, and as a state tool in determining service funding for students (enrolled under five years). See Table 6 for the preliminary cut scores set for winter 2003 for the TEAE. Students who perform at proficient levels on the reading test (level 4) and writing test (level 5) will not count toward state funding for ESL services (Minnesota Department of Education, 2004a). The TEAE is used by the state to determine reclassification of students as fluent English proficient. However, a student's score on the TEAE does not determine whether he or she will receive or continue to receive services (Minnesota Department of Education, 2004b). Individual districts may choose, at their discretion, to offer supportive services for these students based on multiple indicators including their TEAE scores.

The TEAE is given entirely in English and is administered in sections over a series of days. There are three forms of the test: grades 3-4, grades 5-6, and grades 7-8. The 7-8 test is designed to be used for students in grades 7 and above; therefore it is also suitable for those students who enter the system at the secondary level.

Table 6. Preliminary TEAE Cut Scores Set for Winter 2003

		Grade Level Clusters							
		Grades 3-4		Grades 5-6		Grades 7-8		Grades 9-12	
Subject	Level	Scores	% at Level	Scores	% at Level	Scores	% at Level	Scores	% at Level
Reading	4	240	10%	266	12%	278+	8%	278+	12%
	3	187.5-< 240	37%	207-<266	47%	220-<278	50%	230-<286	44%
	2	137.5-187.5	42%	175-<207	26%	180-<220	33%	180-<230	33%
	1	1-<137.5	11%	1-<175	15%	1-<180	9%	1-<180	11%
Writing	5	25+	2%	26+	3%	27+	2%	28+	2%
	4	19.5-<25	31%	22.5-<26	27%	23-<27	26%	23.5-<28	28%
	3	13.5-<19.5	49%	17.5-<22.5	52%	18.5-<23	48%	18.5-<23.5	45%
	2	7.5-<13.5	13%	12-<17.5	12%	16-<18.5	12%	16-<18.5	10%
	1	0-<7.5	7%	0-<12	6%	0-<16	12%	0-<16	16%

Classroom Measures

Classroom measures were collected for three students at each participating school site. These measures included teacher ratings of individual students' skills using a five point Likert scale from 1 (unsatisfactory) to 5 (excellent) in the following areas: reading, writing, listening, and speaking. We asked teachers to use a similar Likert scale to rate a student's (1) ability to master work in class, (2) performance compared to peers, and (3) chance of future success without ESL/bilingual services. ESL/Bilingual teachers and content area teachers also collected 2-3 examples of students' best class work to give support for their rubric-based ratings of students' overall skill levels in reading and writing. These state-created reading and writing rubric descriptors were based on TEAE score levels, and were designed to bridge the TEAE score to identified classroom characteristics of student performance in these areas. See Appendix B for examples of key instruments used with teachers, including background survey, student ability rating scales, and state-developed rubrics. Other materials were provided to teachers, including examples of samples that had been rated during the pilot. These are not included in Appendix B.

State Achievement Tests

MCA

The state assessments of academic achievement were the Minnesota Comprehensive Assessments (MCA) and the Minnesota Basic Skills Test (BST). The MCAs are used for state and federal Title I accountability and are designed to measure student progress toward high content standards in reading, mathematics, and writing (with science to be added in 2006). Until other grades are phased in, the reading test is given in grades 3, 5, 7, and 10. The writing test is given in grade 5 only. Our study used data for tests that corresponded best to the grade level of participants (grade 5 reading and writing, and grade 10 reading scores from spring 2004 for those currently in 11th grade).

BST

Minnesota's BST measures student mastery of minimum competency skills in reading (grade 8), mathematics (grade 8), and writing (grade 10). Students must pass these tests in order to graduate with a standard diploma and the designation of "passed-state level." For this study, we used reading and writing data for students in grade 8, and for students who took the test in later grades (e.g., grade 11).

Procedure

Teachers were instructed to read the list of rubric descriptors for reading and writing, and then to choose two overall levels that reflected the student's skills in reading and writing. Teachers also were asked to collect samples of work and to identify evidence for the level they chose by

highlighting evidence for at least one criterion under each main heading of skills (e.g., vocabulary, comprehension, etc.). If a student had traits that spanned across more than one level, the teacher was still encouraged to choose an overall performance level.

After teachers collected the samples and rated them, they were submitted to project staff. All samples were then rated by two staff members. Each staff member was given the samples in a different order and in separate locations to reduce potential bias. These second ratings were conducted to see whether samples would be rated similarly across multiple reviewers. Where differences occurred, researchers discussed the teacher ratings and evidence. In cases where the teacher and staff differed in ratings, a third staff member was also asked to rate the sample for analysis purposes.

Researchers obtained TEAE, MCA, and BST scores from test databases provided by the state. Researchers used student numbers from the Minnesota Automated Reporting Student System (MARSS) and locally assigned school identification numbers, when needed, to retrieve these achievement scores. The accuracy of one third of the scores recorded in the study database were later checked and found to be accurate at 97%. The one error in this check was then corrected.

Results

This study generated an extensive amount of data, which were analyzed in a variety of ways. These analyses include three parts. First we examined the relationships between the language proficiency measures (e.g., LAS and TEAE) for all participants, and also by smaller groups (e.g., grade) where numbers were sufficient. Second, we compared these results to teacher ratings of classroom samples and addressed issues related to the rating data. Third, we analyzed relationships among language proficiency tests, teacher rating data, and student performance on the state MCA and BST reading and writing test scores. And finally, a special analysis was done for a small subset of students who had scored low on the LAS oral component. This final analysis highlights issues particular to that group in comparing their performance across the other study measures (e.g., ratings, state achievement scores). Initial frequencies by score level were run for each proficiency test. These are provided in Tables C1-C5 in Appendix C.

Connections between Tests of English Proficiency

First it was important to compare the students' results on the two English language proficiency tests used in the study. Correlations were run between the LAS and TEAE (2003-2004). In general, significant correlations were found between the TEAE reading scaled score and the LAS reading score, and between the TEAE writing scaled score and the LAS reading score. Results are presented in Table 7. These data suggest that the underlying reading skills being

measured by the LAS and TEAE are closely related. However, this does not mean that the tests are measuring the exact same skills. It means that the students who performed one way on one test tended to perform a certain way on the other test. The writing tests for the LAS and TEAE were not related at all, indicating that the tests are either measuring different skills, or are measuring underlying skills differently. Because NCLB's requirement of an annual measure of English proficiency alludes to the need for a test that gauges academic English growth, it will be important to compare students' performance on these measures to the other academic measures that are the focus of this study.

Table 7. Correlations between the LAS and TEAE

	LAS-Oral	LAS-Reading	LAS-Writing	TEAE Reading SS[1]	TEAE Writing
LAS-Oral	1.0	.458*	.304*	.309**	.191
LAS-Reading			.321**	.571**	.536**
LAS-Writing				.191	.192
TEAE Reading SS					.415*
TEAE Writing SS					
Total N				96	36

* Correlation is significant at the .05 level (2 tailed).
** Correlation is significant at the .01 level.
[1] SS= Scaled Score

Comparison of Language Proficiency Tests and Work Sample Ratings

In order to help address the question of whether higher performance on these proficiency measures is more related to higher academic achievement in general, we need to focus on comparing the proficiency score results to teachers' ratings of their work and state achievement test results. Before presenting these correlations, it is important to present details about the classroom measures so that correlations can be interpreted with greater accuracy.

Classroom measures included ESL/bilingual and content (science or social studies) teacher ratings of: (1) student abilities in reading, writing, speaking, and listening, (2) students' overall abilities in relation to course content and peers, and (3) student work samples, chosen by teachers to represent the student's best work in reading and writing in his or her class. Although we limited out correlations with proficiency score results to teacher ratings of work samples (i.e., reading and writing), the other teacher ratings about students' skills across modalities and in comparison to their peers are still important for a broader picture of how teachers viewed these students' skills overall.

A total of 25 students had work samples collected and rated. The languages of these students

were: African 4% (N=1), Amharic 4% (N=1), Creolized English 8% (N=2), English 4% (N=1), Hmong 16% (N=4), Russian 8% (N=2), Somali 20% (N=5), Spanish 32% (N=8), Tibetan 4% (N=1).

Teacher Ratings of Student Skills Across Modalities

The results of teachers' ratings of student skills in reading, writing, speaking, and listening are presented in Table 8 (See Tables D1-D4 in Appendix D for the tables used to make Table 8 by grade levels). Many of the samples had only one rating – the rating of either the ESL or content teacher. Those data were removed here to provide a clearer comparison of the same students across type of teacher. In general, ESL teachers rated students the same or higher than content area teachers (see Table D5 in Appendix D for ratings by individual students).

Table 8. ESL and Content Area Teacher Ratings of Skills Across Grades

	Speaking All Grades		Listening All Grades		Reading All Grades		Writing All Grades	
	ESL	Content	ESL	Content	ESL	Content	ESL	Content
5 (Excellent)	6	7	8	6	4	-	3	3
4	8	7	9	4	8	3	6	4
3	5	6	2	5	6	7	7	5
2	-	4	-	4	-	7	2	5
1 (Unsatisfactory)	-	-	-	-	-	1	-	1

Dash indicates no rating for that number on the scale.

Although it is not customary to average the means of Likert scale ratings, we provide Table 9 as a condensed view of the tendency of teachers to rate high or low for specific modalities. Both ESL and content teachers tended to rate Listening highest, followed by Speaking. After these, ESL teachers' next highest rating tended to be Reading (3.66) followed by Writing (3.35). Content teachers tended to have slightly higher ratings for Writing (3.04) than Reading. Overall, ESL teachers had a higher average mean across all ability levels, with their lowest means for reading and writing, roughly equal to the highest average means (Listening and Speaking) for content teachers.

Table 9. Overall Rating Means Across Grades by Teacher and Modality

Skill Area	ESL	Content
Listening	4.14	3.61
Speaking	3.95	3.35
Reading	3.66	2.77
Writing	3.35	3.04

Teacher Rating of Student Skills Related to Academic Work

Students were also rated using a similar five point Likert scale for three additional questions, including student ability to master content, performance in relation to fluent English peers, and student chances of success in future content classes without further ESL support. ESL and content teacher ratings for these questions are presented in Tables 10 and 11. Ratings based on fluent English peers are provided in Appendix Table D6. Most ESL teachers did not have fluent English peers in their classes. Content teachers, for this question, had slight increases in ratings from 5th grade (2s and 3s) to 8th grade (4s and 5s). See Appendix Table D7 for individual student ratings on all three of these additional questions.

In general, in keeping with the pattern established by the modality ratings, ESL teachers appeared to rate students the same or higher on their ability to master content than did content area teachers (see Table 10). In considering students' chance of success without language assistance, there appeared to be a wide range of opinions reflecting the strengths and weaknesses of these students (see Table 11). The responses suggest that these content and ESL teachers varied in their opinions concerning individual student success, with perhaps the most divergence showing among grade 11 teachers. Comparing the two tables, fifth grade ESL teachers tended to rate students' ability to master content and likelihood of success without support as slightly higher than content teachers. Eighth grade ESL teachers tended to rate ability to master content slightly higher than content teachers, but success without support was rated about the same. The 11th grade teachers tended to rate ability to master content about the same, but had very different opinions about the ability of students to succeed without support, some content teachers being more optimistic about certain students than their ESL colleagues, and some being less optimistic.

When we look at the correlations between teacher ratings and proficiency test performance (see page 15), tendencies discussed in this section should be kept in mind: that ESL teachers tended to rate students higher in mastery potential and higher across individual modalities, and that content teachers had more diverging opinions concerning the 11th grade students as to future success without ESL support.

Table 10. ESL and Content Teacher Ratings of Student Ability to Master Content

	5th N=5		8th N=6		11th N=8	
	ESL	Content	ESL	Content	ESL	Content
(Very Capable) **5**	1	1	4	2	5	5
4	1	-	1	2	2	1
3	2	1	1	1	1	2
2	1	3	-	1	-	-
(Unable) **1**	-	-	-	-	-	-

Dash indicates no rating for that number on the scale.
Note: Shading highlights differences.

Table 11. ESL and Content Teacher Ratings of Student Chance of Success without ESL Support

	5th N=5		8th N=6		11th N=8	
	ESL	Content	ESL	Content	ESL	Content
(Excellent) **5**	-	1	2	1	-	3
4	1	-	3	3	3	-
3	3	-	-	-	3	1
2	1	3	1	2	1	3
(Unlikely) **1**	-	1	-	-	1	1

Dash indicates no rating for that number on the scale.
Note: Shading highlights differences.

Teacher Ratings of Work Samples

Teachers also rated student work samples using state-developed descriptors for reading and writing. These descriptions are based on the TEAE. Teacher ratings using the rubric descriptors reflected a broader range of students' skills in the fact that teachers made full use of the 1 to 5 rubric, especially in writing (see Appendix Tables D8 and D9). In Tables 12 and 13, we show the rubric ratings for only those students with ratings by both the ESL and content teacher because we can make comparisons of the ratings for these students. These tables indicate that content teachers for the grade 8 and 11 students tended to rate skills as similar to or lower than the ESL teachers. This is in contrast to the students in grade 5, where the content teachers rated their skills slightly higher than their ESL teachers in both reading and writing.

Table 12. Teachers Ratings of Reading Samples Using Rubrics

Reading Rubric Descriptor Level	5th N=4		8th N=3		11th N=5	
	ESL	Content	ESL	Content	ESL	Content
5	-	-	-	-	-	-
4	-	2	1	-	2	2
3	3	2	2	2	3	3
2	1	-	-	1	-	-
1	-	-	-	-	-	-

Dash indicates no rating for that number on the scale.

Project staff ratings of student work samples were compared to those of the ESL and content area teachers. The percent agreement between the two project staff raters and ESL teachers was 78% for each rater. Agreement among the two staff raters ranged from 71-100%. Agreement between content teachers and staff ratings ranged from 54%-63% for those same samples. Among staff raters this agreement was 56%-69%. Although some of these agreement rates were small, they are often due to incomplete samples submitted by teachers (e.g., teachers provided answers to questions but the reading passage and the questions were not provided, or teaches provided

Table 13. Teachers Ratings of Writing Samples Using Rubrics

Writing Rubric Descriptor Level	**5th** N=4		**8th** N=5		**11th** N=8	
	ESL	Content	ESL	Content	ESL	Content
5	-	1	-	-	1	1
4	-	-	4	3	4	2
3	4	2	-	2	3	4
2	-	1	1	-	-	1
1	-	-	-	-	-	-

Dash indicates no rating for that number on the scale.

student answers that looked like reproductions of a text). These issues led to somewhat different approaches to rating the samples. Details of the ratings are provided in Appendix Table D10. In general, higher agreement ratings were found between project raters than between the teachers and the project raters. Higher rates of agreement were also found when using work samples provided by the ESL teacher rather than the content area teacher. These results may be partly due to the fact that one of the project staff raters, and a staff rater (rater 3) that provided additional input on difficult samples (though not counted in final ratings), had ESL backgrounds. This and other differences in ESL and content teacher ratings are addressed in the Discussion.

Correlations of Language Proficiency Tests and Work Sample Ratings

The relationships between English language proficiency test performance and the ratings of students' reading and writing work samples were examined through correlations. Tendencies discussed in the previous section, specifically, the generally higher ratings of ESL teachers, should be kept in mind in interpreting the correlation results. In order to maintain the differences in ESL and content area teachers' ratings of student samples, correlations were run by teacher type.

The relationship between teacher ratings and proficiency scores on the LAS and TEAE generally were very different (see Table 14). Significant relationships were found between teachers' ratings of work samples and TEAE scores, especially between ESL teacher ratings of writing samples and TEAE scores in reading and writing. Content teachers also showed a significant correlation, though to a lesser degree, between writing sample ratings and the TEAE reading score. The LAS scores did not show as significant a relationship to either ESL or content teachers' ratings of reading or writing. ESL writing sample ratings were modestly correlated to the LAS writing score. Other correlations within teacher ratings (e.g., ESL reading correlates to ESL writing scores) are also shown. Caution should be exercised in interpreting some of the correlations because of the small numbers of students (numbers are included in parentheses).

Table 14. Correlations between Language Proficiency Tests and Teacher Ratings

	Sample **Reading** ESL Teacher		Sample **Writing** ESL Teacher		Sample **Reading** Content Teacher		Sample **Writing** Content Teacher	
	Corr.	N	Corr.	N	Corr.	N	Corr.	N
LAS Oral	.006	*(17)*	.202	*(20)*	-.021	*(19)*	-.223	*(20)*
LAS Reading	.179	*(17)*	.426	*(20)*	-.117	*(19)*	-.079	*(20)*
LAS Writing	-	*(17)*	.548*	*(20)*	-	*(19)*	-	*(20)*
TEAE Read Scaled Scr	.541*	*(17)*	.602**	*(19)*	.231	*(19)*	.595**	*(20)*
TEAE Write	.592	*(7)*	.955**	*(8)*	.725	*(7)*	.470	*(9)*
Sample Read ESL	1.0	*(17)*	.699**	*(17)*	.365	*(13)*	.393	*(14)*
Sample Write ESL	.699**	*(17)*	1.0	*(20)*	.268	*(15)*	.408	*(16)*
Sample Read CON	.365	*(13)*	.268	*(15)*	1.0	*(19)*	.613**	*(18)*
Sample Write CON	.393	*(14)*	.408	*(16)*	.613**	*(18)*	1.0	*(20)*

Dash indicates no rating for that number on the scale.
* Correlation is significant at the .05 level (2 tailed).
** Correlation is significant at the .01 level.
Shading highlights significant relationships between TEAE and LAS scores to sample ratings.

Connecting Proficiency Tests, Classroom Ratings, and Standardized Tests

The potential relationships between classroom performance and student achievement tests were investigated by comparing the overall mean reading and writing sample ratings of ESL teachers and content teachers to the means of student performance on the TEAE, the MCA, and the BST where these scores were available (see Table 15). For tables used to create Table 15, see Appendix Tables D11 and D12. The TEAE was used here alone because we wanted to see specifically whether teacher ratings using student descriptors based on the TEAE would correspond well in this sample.

Teacher Work Sample Ratings Compared to Proficiency Assessments and Achievement Tests

In general, average test scores appeared to follow the same pattern as teacher ratings. As ratings increase, the average scores on the TEAE, MCA, and BST also increase. All average MCA reading scores were rated at least at level IIb, which indicates students are meeting grade level expectations. This means that students who were rated as a 3 or 4 in reading by their ESL teachers had an average BST reading score that allowed them to pass the test. Content teacher ratings of students' reading skills did not relate strongly to their scores on standardized reading tests. It is notable that two students who had been rated a "2" by a content teacher achieved passing scores on the BST (see Appendix Table D12). It should also be noted that some students may have passed but were in a group where the mean was "Not Passing," as in the group rated level 3 by content teachers but which had an overall group mean of 594.5 (see Appendix Table D12).

The small numbers of students represented in the data requires that caution be exercised in inter-

preting findings. Writing is not presented here because there was not enough state performance data for the students in this study to be meaningfully represented here.

Table 15. ESL and Content Teacher Reading Sample Rating Means by TEAE, MCA and BST Mean Performance

	ESL Teacher Rating				Content Teacher Rating			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
TEAE R level (N)	-	2 (4)	3 (10)	High 3 (3)	3 (1)	3 (5)	3 (8)	3 (5)
MCA R level (N)	-	IIB (2)	IIB (3)	-	IIA (1)	IIA (1)	IIB (2)	IIB (3)
BST R (N)	-	Mean Not Passing (2)	Mean Passing (5)	Mean Passing (3)	-	Mean Passing (2)	Mean Not Passing (4)	Mean Passing (2)

Dash indicates no rating for that number on the scale.

Students Rated Likely to Succeed Compared to Language Assessment and State Achievement Test Data

Table 16 presents the data for students who were rated by their teachers as having a near excellent (4) to excellent (5) chance of succeeding in future classes without additional help in learning English. Students were included here if either teacher rated their chances as excellent or near excellent (4 or 5).

Generally, students who had scored at least a 3 on TEAE Reading and a 4 or higher on TEAE writing tended to have at least one teacher who gave them an excellent rating. For example, among fifth graders, student 1 performed at level 4 (proficient by state determination) and also performed at the second highest level on the reading and writing MCA tests (IV). Yet, the ESL teacher rated the student as having an average chance of success in future classes without language support in contrast to the 5 given by the content teacher. Student 2, also a fifth grader, although performing at the highest level on the MCA writing test, did not test at the proficient level on the TEAE writing test.

For eighth graders, four out of five students had passed the BST reading test. All of the 4 students' teachers gave their student a 4 or 5, except one content teacher who gave a 2 rating.

For grade 11 students, two of five students who received at least one 4 or 5 had passed the BST reading test. But like the grade 8 students, there was one content teacher who gave a lower rating (2). For students 9 and 10, the data do not fit a pattern. The two content teachers gave very high ratings, in contrast to the ESL teachers' average ratings. Because these two students had low TEAE reading scores, it is uncertain as to why the content teachers had been so assured in the

students' abilities to succeed without further language assistance. It may be that these teachers were answering the question in a more future oriented mindset or were basing their response on other student characteristics such as determination. In comparison to state ELL data, the students in this table show above average scores for the MCA reading test (1378.74) and BST (585.50) for school year 2002-03 (Kato, Albus, Liu, Guven, Thurlow, 2004).

Table 16. Scores of Students Rated Likely to Succeed without Additional Language Assistance

Student		Rating	TEAE Reading		TEAE Writing		MCA Reading		MCA Writing	BST Reading
Grade	Teacher		SS*	L*	S	L	SS	L	L	
1	ESL	3	288	4	-		1720	IV	1840 IV	
5th	CON	5								
2	ESL	-	232	3	26	4	1470	IIb	1970 V	
5th	CON	5								
3	ESL	4	222	3	23	4	1550	III	1690 IV	
5th	CON	2								
4	ESL	5	247	3	26	4				600
8th	CON	5								
5	ESL	5	260	3	-					617
8th	CON	2								
6	ESL	4	230	3	26	4				566
8th	CON	5								
7	ESL	4	265	3	-					611
8th	CON	4								
8	ESL	4	303	4	-					617
8th	CON	4								
9	ESL	3	211	2	-					584
11th	CON	5								
10	ESL	3	200	2	-					-
11th	CON	5								
11	ESL	4	303	4	29	5				654
11th	CON	5								
12	ESL	4	233	3	-					617
11th	CON	2								
13	ESL	4	251	3	-					-
11th	CON	3								

Dash Indicates no rating or no score found.
* S=Score, SS= Scaled score, L = Level

Correlations for Language Proficiency Tests, Sample Ratings, and State Achievement Tests

Several significant relationships were found in the comparison of language proficiency tests, sample ratings, and state achievement test scores (see Table 17). Although certain LAS scores (reading and overall) were significantly correlated to the MCA writing test and BST reading test, there appear to be stronger correlations between the TEAE reading score and the MCA and BST reading scores. A modest correlation was found between the TEAE reading and MCA

writing test. The TEAE writing score had a strong correlation with BST reading, though the N for this correlation is small (17). ESL reading and writing samples also were correlated with the BST reading test, however these numbers are also small (10 and 11).

Table 17. Correlations between Language Proficiency Tests, Sample Ratings, and the MCA and BST

	MCA Reading		MCA Writing		BST Reading	
	Corr	N	Corr	N	Corr	N
LAS Level	.408*	27	.561**	28	-.126	49
LASO	.183	27	.222	28	.116	49
LASR	.220	27	.224	28	.513**	49
LASW	-.082	27	.257	28	.111	49
TEAE R SS	.693**	27	.433*	28	.701**	49
TEAE W	.349	16	.445	16	.805**	17
Sample Read ESL	.289	5	.474	5	.738*	10
Sample Write ESL	.175	6	.351	6	.772**	11
Sample Read CON	.291	7	.579	7	.071	8
Sample Write CON	.670	6	.513	6	.341	10

* Correlation is significant at the .05 level (2 tailed).
** Correlation is significant at the .01
Shading highlights significant relationships.

Data for ELLs with Very Low LAS Oral Scores

Students with very low overall oral scores on the LAS generally would not be given the opportunity to take the reading and writing portions of the LAS (see LAS scoring process). In this study, however, we allowed these students to take both the reading and the writing LAS tests.

LAS Scoring Process

The LAS scoring process assigns a final LAS score that indicates students' overall proficiency in English. This score uses all parts of the reading, writing, listening and speaking tests. The code of LEPa is assigned to students with low reading and writing scores, but mid level listening and speaking scores. LEPb includes students with low reading and writing scores, but high listening and speaking skills. LEPc includes students with mid level reading and writing scores as well as mid level listening and speaking skills. LEPd is assigned to students who display mid level reading and writing skills and high listening and speaking skills. LEPe includes students

with high reading and writing skills and mid level listening and speaking skills. Fluent English Proficiency (FEP) is used to designate students with high level reading and writing skills and a high level of listening and speaking skills. It should be noted that if a student has a score ratio of 1/1, 2/1, or 3/1, that student is not included because the LAS authors "do not recommend administering the LAS R/W to students whose oral proficiency is lower than the equivalent of LAS-O Level 2." Students with this type of score are included in the "N/A" column. The frequencies of overall LAS student scores can be found in Table 18. In general, the majority of our sample included students in all of the language proficiency categories with the exception of LEPb. The majority of students appeared to demonstrate mid level proficiency, particularly in the later grades.

Table 18. Frequencies for Overall LAS Score

Grade	LEPa Low R/W Mid L/S	LEPb Low R/W High L/S	LEPc Mid R/W Mid L/S	LEPd Mid R/W High L/S	LEPe High R/W Mid L/S	FEP* High R/W High L/S	N/A[1]
5th	2	0	6	13	2	3	2
8th	2	0	19	4	4	1	1
11th	3	0	19	1	3	1	5
Total	7	0	44	18	9	5	8

* FEP= Fluent English proficient

Low Oral Score Group Performance and Ratings

Table 19 provides the reading and writing performance of students whose oral proficiency is lower than the equivalent of LAS-O level 2 (i.e., the 8 students in the N/A column in Table 18). All students in Table 19 received a level "1" score for the oral and listening component of the LAS. We note that the majority of these students (like the others in our study) performed at Level 2 for writing, and we draw attention to the fact that Student 7 had the highest possible score in reading.

Although teachers had spoken to staff about "attending" to approximately three students during the course of the LAS administrations (e.g., help keep student on track), only one study participant (Student 1 in Table 21) was formally identified in the state database as receiving special education services. We emphasize here that this student also was able to participate in the reading and writing tests even though the resulting score was in fact a "1." The possibility that other students may have a disability that directly relates to oral and listening skills should not prevent them from showing what they know in another modality which may actually be a student's strength.

We include TEAE levels for the eight students alongside the LAS score levels in Table 19 for comparison purposes. Student 8, who had scored the highest level on LAS reading, performed

just at level 2 on the TEAE (on a four level scale). However, two other students performed at level 3 on the TEAE writing test (on a five level scale), indicating mid-level skills in writing. The individual student skills in reading and writing reflected here would not be known in the recommended administration of the LAS test based on the overall oral score which includes listening components.

Table 19. Students with Lowest LAS Oral Score with Reading and Writing Levels from LAS and TEAE

	Grade	LAS Oral	LAS Reading	LAS Writing	TEAE Reading	TEAE Writing
Student 1	5	1	1	1	1	*
Student 2	11	1	1	2	2	*
Student 3	8	1	1	2	2	2
Student 4	11	1	1	2	1	2
Student 5	11	1	2	2	1	3
Student 6	11	1	2	2	2	3
Student 7	5	1	1	1	*	*
Student 8	11	1	3	2	2	*

*Indicates missing score.

Of these eight students, two had reading and writing samples collected and rated by teachers (see Table 20). The data show another perspective of these student skills. The ESL teachers' ratings indicate low to medium level skills in reading and writing in comparison to slightly higher ratings by content teachers.

Table 20. Teacher Ratings for Students in Work Sample Group with Very Low Oral LAS Scores

	Grade	LAS Oral	LAS Reading	LAS Writing	Reading Sample ESL Teacher	Reading Sample Content Teacher	Writing Sample ESL Teacher	Writing Sample Content Teacher
Student 3	8	1	2	2	2	*	2	4
Student 8	11	1	1	1	*	3	*	3

*Indicates missing rating.

Table 21 shows BST performance for five of the students who had the low overall LAS oral and listening score. A score of 600 is required to pass the basic skills reading test. In comparison to state data for ELLs, although these students are below the state mean for reading (585.50), a few are almost on the boundary of the mean range (SD = 44) for school year 2002-2003 (Kato et al., in press).

Table 21. Performance on State Tests for Group with Very Low Oral LAS Scores

	Grade	LAS Oral	LAS Reading	LAS Writing	BST Reading Score
Student 2	11	1	1	2	558
Student 3	8	1	1	2	518
Student 4	11	1	1	2	545
Student 5	11	1	2	2	532
Student 6	11	1	2	2	554

Discussion

As we quoted at the beginning, "little research exists on the relationship between academic language, language proficiency tests, and performance on standardized content assessments" (Stevens, 2000, p. 4). This study sought to provide more information on these links by examining relationships between the language proficiency measures (e.g., LAS and TEAE), teacher ratings of classroom samples to the proficiency measures, and then finally each of these to Minnesota's MCA and BST reading and writing test scores.

In examining the relationship between student performance on the TEAE and LAS, we found that the underlying reading skills being measured by the LAS and TEAE were closely related. This does not indicate that the tests are, however, measuring the same skills. Rather it indicates only that the students who performed one way on a test tended to perform a certain way on the other test. The writing tests for the LAS and TEAE were not related indicating that the tests are either measuring different skills, or are measuring underlying skills differently.

We also collected teacher ratings of student achievement in speaking, listening, reading, and writing. When we compared the ESL teacher ratings to those of the same student's content teacher, we found that ESL teachers tended to rate the student's skills as either the same or higher than the student's content teacher. In general, both ESL and content teachers tended to rate listening and speaking skills higher than the students' skills in reading or writing. When asked about students' ability to succeed without further language support, there was a wide range of opinions. When this was further explored, most of the students rated as likely to succeed had achieved passing scores on the BST and had at least been rated as "achieved" on the MCA tests, though sometimes a student with lower teacher ratings was also achieving passing scores on state measures such as the BST Reading test.

A comparison of all three measures (i.e., language proficiency scores, teacher ratings, and standardized statewide achievement assessments) showed that certain LAS scores (reading and overall) were significantly correlated to the MCA writing test and BST reading test. However,

there appear to be stronger correlations between the TEAE reading score and the MCA and BST reading scores. A modest correlation was found between the TEAE reading and MCA writing test. The TEAE writing score had a strong correlation with BST reading, though the N for this correlation was small (17). ESL reading and writing samples were also correlated with the BST reading test, however these numbers are also small (10 and 11).

Comparisons of ESL teacher ratings of student ability to state achievement tests showed that as teacher ratings increased, average scores on standardized tests also increased. However, when comparing content teacher ratings of students' skills to standardized tests, there was no clear pattern. Some students who were rated as having low reading skills had actually passed the BST. In general, content teachers' ratings of student writing skills were more consistent with standardized test scores. Some students rated lower by their teachers actually achieved higher average scores on the TEAE writing test than other students who had been rated more highly in that skill area.

In general, content teacher perceptions of student skills, as indicated by ratings were different from ESL teachers' perceptions. This may be due to general background differences between ESL and content teachers who may have varied in their ability to shift from evaluating student language use versus content performance in class as interpreted through the rubrics. ESL teachers are trained to focus on language development. Thus, it is entirely possible that this perspective may have contributed to their more uniform results with each other (e.g., work samples using rubric linked to TEAE) in comparison to state measures (e.g., TEAE) based on similar levels of language development.

Finally, for the group of students with low LAS oral scores, we sought to address the need to reevaluate assumptions about the order of demonstration of skill acquisition in learning other languages, and the possibility that a student may have a disability leading him or her to favor one modality of expression over another in the student's first or second language. We found that giving these students the opportunity to take the reading and writing sections of the test allowed them to show a range of skills in these other areas. Although many of these students did have low reading and writing scores as shown on the LAS and TEAE, some of them had scores comparable to their peers on these same tests. Further, for those students who did have low levels of reading and writing, at least they were given the opportunity to take the reading and writing assessment. This showed (1) a willingness to accept the fact that individual students, whether by their own pattern of development or possible disability, may vary in their language proficiency across modalities, and (2) that value is placed on giving students the opportunity to show what they are capable of rather than having this information assumed for them. Although these same students' scores on state tests showed below average skills compared to previously examined state data, these findings still underscore the need for simultaneously administered

measures of language proficiency across modalities and care in using results of assessments for instructional decision-making.

This study was conducted to examine relationships among measures of English proficiency, standardized content assessments, and classroom performance. By examining these, we begin to understand the complexity and nuances of how teachers view student performance from language and content perspectives and how these in turn relate to language acquisition and academic achievement. By addressing these relationships, several assumptions have been made in the administration of certain assessments for students learning a second or other language.

Difficulties Conducting the Study

It was very challenging to recruit districts and schools to participate in this study. Over 20 districts were contacted for possible participation. Eight of these did not return repeated phone calls. Of the districts that verbally declined to participate in the study, all of them cited over-testing of students and loss of instructional time as reasons for their refusal. Additional reasons included: lack of translated consent forms for languages of students who were not the focus of the study (Hmong, Spanish, and Somali language background students were the focus of the study). Although English consent forms were provided for all students, translation costs prohibited the translation of the study forms into languages other than the three targeted language groups.

Other circumstantial issues that made teacher participation difficult included: stressed schedules, localized friction, frustrations with past state departmental decisions, failure to pass local funding referendums, a major district's decision to close a number of schools, and tensions over uncertain layoffs. Clearly, efforts to conduct research that involves additional testing are likely to be unsuccessful or unwelcome unless these efforts have the full support of a broad section of teachers and state level promotion for participation.

During the time frame of this study, information was publicly released suggesting future changes to the LAS, and potential issues with the Illinois' Measure of Academic Growth in English (IMAGE), the test on which the TEAE was based. The publishers of the LAS test (Jackson et al., 2003) indicated that a new LAS would be better aligned to standards in response to Title III. In addition, a Chicago Tribune article (Dell'Angela & Cohen, 2004), suggested that changes might be underway for the IMAGE test either by replacing it with a new assessment by 2006 or enhancing the existing test more oriented to reading achievement. This type of revision has not been suggested by Minnesota.

Limitations

Although finding adequate numbers of subjects and teachers willing to participate was problem-

atic and limited our ability to represent schools outside the state metropolitan area, there were additional limitations to the interpretation of the data. These include the function of student familiarity with the LAS test in certain districts, the delay in collecting teacher rating samples, and other related issues.

Concerning the issue of student familiarity, one suburban district and one urban district had used the oral LAS test one to two times a year. Although the LAS test that was used for this study, especially the oral component, was selected because it was thought that districts did not use it, some students communicated to the researchers that they were familiar with the oral component. This was particularly true in one suburban district where students had taken it the previous spring. This suggests that there is a need for caution concerning the interpretation of a few of the higher oral scores for fifth grade students from that district.

Another limitation is the small number of students in many of the correlations. The underlying issue with small numbers was the lack of data for many of the students, including the dearth of writing scores in the state database and the difficulty in getting an ESL and content area teacher to complete work sample ratings. Also, because this study straddled several grades (originally designed for 150 students, with 50 in each grade), the amount of comparison data across statewide tests was less than what it could have been if the study had focused on one grade level.

Caution also should be used in interpreting the teacher samples. Teachers varied considerably in the time they took to complete this portion of the study (ranging from 2 weeks to 6 months). Although the majority of teachers finished the collection and rating of work samples within 2 to 4 weeks of testing, there were a few teachers who had longer delays for providing a rating of student work. Also, for sites secured later in the year, the TEAE scores had already been distributed. It is not known whether or to what extent the release of TEAE scores to the schools may have influenced teacher ratings of student work. Teachers do not regularly use the descriptors in instructional settings, so the linkage between the TEAE test scores and descriptor rubrics may have had little or no effect on the teacher ratings. In support of this conclusion we point out that staff ratings which were done without knowledge of student TEAE performance or teachers' ratings were more in line with ESL teachers' ratings, suggesting that perhaps the similarity is more likely due to approaching the task from an ESL perspective. One of the project staff raters, and an additional staff rater that provided input on difficult samples (see Tables D8-D10 in Appendix D), had ESL backgrounds.

This and other potential issues may be underlying the variability in teachers' ratings. It is possible that teachers adjusted ratings based on their own evaluation of the difficulty of their class. On the other hand, some students may perform better on class work than they do on paper and pencil tests. Further, there are potential questions about the alignment of instructional tasks with the rubric used to measure student skills that have already been alluded to in the explana-

tions offered for the differences in content and ESL teacher ratings. For example, some content teachers did not easily find examples of writing in their classes that would capture the essay-like characteristics described in the rubric (e.g., writing in paragraphs). Many of the content teachers' samples showed one paragraph or short answer formats in response to readings in a content area. It is uncertain to what extent this type of mismatch in format may have influenced the ability of the rubric to capture student language abilities in academic writing, especially in content area classes.

Finally, student classwork samples, as intended in the study design, were not standardized. This study focused on English language learners in Minnesota and their ability to do school work in non-laboratory day to day classes. Even though staff and teachers rated samples with the same rubric, questions arise similar to those found in portfolio studies (Koretz, Stetcher, Klein, & McCaffrey,1994). Did teachers really choose the best sample to illustrate student skills? Did the rubric reflect the skills students were learning?

We recognize these issues, and offer the findings in the context of how they relate to other more objective measures of language proficiency and achievement. Also, we recognize the complexity of conversing about these results in front of a larger backdrop of different conceptualizations of what academic language is, as outlined at the beginning of this report. The assessments in this study focused on only two proficiency measures and two state achievement tests in reading and writing. More is to be learned from studying other language proficiency measures that have different underlying concepts of academic language and how it is measured. Further study by states comparing these proficiency measures with student progress on newly established state ELD standards and existing state academic content standards will no doubt be valuable in furthering the field's understanding of how English language learners' growth in language proficiency relates to growth and achievement in academic language and content.

References

Abedi, J., Leon, S., & Mirocha, J. (2001). *Examining ELL and non-ELL student performance differences and their relationship to background factors: Continued analyses of extant data.* Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing

Chamot, A. U., & O'Malley, J. M. (1994). *The CALLA handbook: How to implement the Cognitive Academic Language Learning Approach*. Reading, MA: Addison-Wesley.

Collier, V.P. (1992). A synthesis of studies examining long-term language minority student data on academic achievement. *Bilingual Education Research Journal, 16*(1/2), 187221.

Cummins, J. (1984). *Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students*. In C. Rivera (Ed.), *Language proficiency and academic achievement*. Clevedon: Multingual Matters.

Cunningham, J.W., & Moore, D.W. (1993). The contribution of understanding academic vocabulary to answering comprehension questions. *Journal of Reading Behavior, 25*(2), 171-180.

De Avila, E. A., & Duncan, S. E. (1990). *Language assessment scales: Oral technical report.* Monterey, CA: CTB/McGraw-Hill.

Dell'Angela, T., & Cohen, J. S. (2004, February 20). U.S. eases test rules for immigrant kids. *Chicago Tribune*. Retrieved February 20, 2004, from http://www.chicagotribune.com

DelVecchio, A., & Guerrero, M. (1995, December). *Handbook of English language proficiency tests*. Albuquerque, NM: Evaluation Assistance Center-Western Region, New Mexico Highlands University.

Jackson, S. L., & Jackson, L.G. (2003, January). *Past, present, and future LAS results for program evaluation and student progress.* Paper presented at the meeting of the National Association on Bilingual Education, New Orleans, LA.

Kato, K., Albus, D., Liu, K., Guven, K., & Thurlow, M. (2004). *Relationships between a statewide language proficiency test and academic achievement assessments* (LEP Projects Report 4). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Koretz, D., Stecher, B., Klein, S., & McCaffrey, D. (1994). The Vermont portfolio assessment program: Findings and implications. *Educational Measurement: Issues and Practice*, 13 (3), 5-16.

Minnesota Department of Education (2004a). *ELL education program guidelines: 2004 edition. Funding.* Retrieved August 4, 2004 from http://education.state.mn.us/content/056347.pdf

Minnesota Department of Education (2004b). *ELL education program guidelines: 2004 edition. Procedure.* Retrieved August 4, 2004 from http://education.state.mn.us/content/056347.pdf

No Child Left Behind Act of 2001, 20 U.S.C. 6301 et seq. (2002).

Olson, L. (2002, December 4). States scramble to rewrite language-proficiency exams. *Education Week.* 22 (14),10.

Ramírez, J., Yuen, S., Ramey, D., & Billings, D. (1991). *Final report: Longitudinal study of structured English immersion strategy, early-exit and late-exit bilingual education programs for language-minority children (Vols. I, II)* (No. 300-87-0156). San Mateo, CA: Aguirre International.

Scarcella, R. (2003). *Academic English: A conceptual framework*. Irvine, CA: University of California, Linguistic Minority Research Institute.

Solomon, J., & Rhodes, N. (1995). *Conceptualizing academic language* (Research Rep. No. 15). Santa Cruz: University of California, National Center for Research on Cultural Diversity and Second Language Learning.

Stevens, R. A., Butler, F. A., & Castellon-Wellington, M. (2000). *Academic language and content assessment: Measuring the progress of English Language Learners*. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

United States Department of Education (2003). *Non-regulatory guidance on the Title III state formula grant program, Part II: Standards, assessments, and accountability*. Washington, D.C.: Author. Retrieved February 25, 2003 from http://www.ed.gov/programs/nfdp/NRG1.2.25.03.doc

Appendix A

Information Provided by Teachers

Table A1. Information Provided by ESL Teachers

District	Grades	Number of ESL Teachers	Areas of Licensure	Licensure Grades	Number of Years Teaching	Number of Years in District
Urban	K-8	3	ESL, Spanish, Elementary Ed (2) Elementary Ed (1)	K-5, 6-8	4-7 years (2) More than 7 years (1)	More than 7 years (3)
Urban	9-12	1	ESL K-12	9-12	1-3 (ESL 6-8), 4-7 (ESL 9-12)	4-7 years
Suburban 1	K-5	1	ESL K-12, Elementary Ed 1-6	K-5	1-3 years	1-3 years
Suburban 1	6-8	1	Elementary Education, English Language Arts for grades 6-8, K-12 ESL	6-8	4-7 years	4-7 years
Suburban 1	9-12	4	ESL, Russian, German (1) ESL (1) ESL K-12, Spanish K-12 (2)	9-12	Less than 1 year (1) 1-3 years (1) 4-7 years (1) More than 7 years (1)	Less than 1 year (2) 1-3 years (2)
Suburban 2	5-6	2	K-12 ESL, Secondary language arts, Spanish, secondary reading endorsement (1) ESL, Elementary Education (1)	K-5, 6-8	4-7 years (2)	4-7 years (2)
Suburban 2	10-12	2	ESL, German (1) ESL K-12, Spanish K-12 (1)	9-12	1-3 years (1) More than 7 years (1)	Less than 1 year (1) 4-7 years (1)

Table A2. Information Provided by Content Area Teachers

District	Grades	Number of Content Teachers	Areas of Licensure	Licensure Grades	Number of Years Teaching	Number of Years in District
Urban	K-8	3	Elementary Education (1) 1st through 6th education, Media Specialist (1); Science (1)	K-8	More than 7 years (3)	More than 7 years (3)

Table A2. Information Provided by Content Area Teachers (continued)

District	Grades	Number of Content Teachers	Areas of Licensure	Licensure Grades	Number of Years Teaching	Number of Years in District
Urban	9-12	1	7-12 Social Studies	9-12	More than 7 years	More than 7 years
Suburban 1	K-5	2	elementary Ed, grades 1-6 (1) Math, Lang Arts, Health, Science, Social Studies, EBD, LD, Elementary Ed. (1)	K-5	More than 7 years (2)	More than 7 years (2)
Suburban 1	6-8	1	Elementary Education K-6, 5-8 middle social studies	6-8	1-3 years	1-3 years
Suburban 1	9-12	2	Social Studies secondary (2)	9-12	More than 7 years (2)	1-3 years (1) More than 7 years (1)
Suburban 2	5-6	5	Elementary Education/ Middle School Mathematics (1) Elementary Education (K-6), Middle School Math (6-8) (1) Social Studies, Family and Consumer Science (1) Grades 1 through 5 (1) 1-6 Elementary Education, Professional Admin-Elem. Principal (1)	K-5, 6-8	Less than 1 year (2) 4-7 years (2) More than 7 years (1)	Less than 1 year (2) 4-7 years (1) More than 7 years (2)
Suburban 2	10-12	3	K-12 Art (1) Social Studies (7-12)	9-12	1-3 years (2) 4-7 (US pol), 7+ (US his)	1-3 years (2) 4-7 years (1)

Appendix B

Instruments

Teacher background survey:

1. What job title do you currently have? (Check all that apply)

____ Content Area Teacher
____ ESL/ Bilingual ed. Teacher
____ LEP supervisor/coordinator/director
____ School administrator
____ Other: please describe _____

2. Content Teacher

For what subject(s)?	For which grades?	For how long?	
_____	___K-5 ___ 6-8 ___9-12	___Less than one year ___1-3 years ___4-7 years ___More than 7 year	
_____	___K-5 ___ 6-8 ___9-12	___Less than one year ___1-3 years ___4-7 years ___More than 7 year	
_____	___K-5 ___ 6-8 ___9-12	___Less than one year ___1-3 years ___4-7 years ___More than 7 year	(list)

3. ESL / Bilingual Teacher. If a bilingual teacher, please mark a **B** next to the subject taught.

For what subject(s)?	For which grades?	For how long?	
_____	___K-5 ___6-8 ___9-12	___Less than one year ___1-3 years ___4-7 years ___More than 7 years	
_____	___K-5 ___6-8 ___9-12	___Less than one year ___1-3 years ___4-7 years ___More than 7 years	
_____	___K-5 ___6-8 ___9-12	___Less than one year ___1-3 years ___4-7 years ___More than 7 years	(list)

4. What areas are you licensed to teach? _____

5. How long have you been teaching in your current district?
(check one) ___Less than one year ___1-3 years ___4-7 years ___More than 7 years

6. Regarding the LAS: When was the last time the LAS was given? _____ What part(s) were administered? (please circle)
Oral Listening Reading Writing

Student_____ Grade _____ (student gender) M___ F___ School _____ Date_____
Teacher _____ (check one) ELA____ Social Studies ____ (teacher gender) M___F___

Student skills:
Please rate the skills of this student in relation to adequacy for doing well in your class. The section on the right is for any comments.

							Comments?
Speaking skills:	1 unsatisfactory	2	3 average	4	5 excellent		
Listening skills:	1 unsatisfactory	2	3 average	4	5 excellent		
Reading skills:	1 unsatisfactory	2	3 average	4	5 excellent		
Writing skills:	1 unsatisfactory	2	3 average	4	5 excellent		

Ability of student to master content in your class:	1 unable	2	3 average	4	5 very capable		
Performance compared to English speaking peers reflected by grades given:	1 unsatisfactory	2	3 average	4	5 excellent	NA*	
Chances of success in regular classes with no additional help in learning English:	1 unlikely	2	3 average	4	5 excellent		

Instructions for Sample Collection

Step 1: Look at the attached list of descriptors for reading and writing skills. Please assess this student using these descriptors and determine what level you think characterizes this student's **best** classroom work. (**circle the level heading on that sheet**.) If you notice that other descriptors on either side of the level you chose also fit the student, please note this on the sheet but you do not need to illustrate these.

Step 2: Next we ask that you provide classroom samples of student work that demonstrate the specific descriptors in that level. All work must have been done by the student by his or herself, with minimal support from the teacher or peers.

Each descriptor category (e.g., vocabulary) has an assigned highlighting color, and specific descriptors are numbered under each category. In the work samples that you collect, we ask that you highlight and code the places that demonstrate a descriptor. There may be more than one color and/or code used per work sample. You do not need to color and code a numbered descriptor multiple times. (Descriptors coded in a certain color & number will be matched to the level heading you chose in step one.)

NOTE: We realize that some descriptors are easier to illustrate than others. We encourage you, after you assess the general level of the student, to provide as many descriptor examples as you can to support your level choice, including those that may not be as easy to show support for. A reasonable number may be one descriptor for each subcategory (e.g., vocabulary, structures, comprehension, etc.)

Reading: Specific Instructions

To illustrate a reading descriptor you have several options: student may give a written, oral, and/or multiple choice response to something he or she has read. Please include a copy of the portion of reading material needed to illustrate meeting a descriptor, as well as a copy of the reading instructions/prompts/questions. A tape is provided for you if you choose to include oral responses. If a tape is used, simply note on a separate paper what descriptor is being demonstrated and describe briefly what part of the student's response matches the specific descriptor(s).

Writing: Specific Instructions

To illustrate a writing descriptor, please include actual writing samples with a description of the writing task/prompt. These samples may be handwritten or word processed. Samples of student writing do not have to demonstrate descriptors across genres, however, to demonstrate certain descriptors in higher grades the inclusion of different genres of writing is required (e.g., grade 9-12 "***Follows prompt genre***").

Example of Reading Descriptors for Grade 11

(These levels 1-4 were converted to one landscape 8" x 17" sheet for ease of use with teachers)

Level 1	Level 2
Vocabulary 1. Very limited sight word knowledge 2. Emerging understanding of sound symbol relationship 3. Limited to basic, concrete vocabulary **Structures** 1. Focus is at the word and/or simple sentence level **Comprehension** 1. May decode without comprehension 2. Needs graphic or picture support to establish meaning 3. Does not connect ideas within text 4. Does not make inferences 5. Can match some words together **Strategies** 1. Reads word by word 2. Lacks strategies for decoding and understanding unfamiliar words 3. Scans at the word level **Texts** 1. Requires graphics to provide context and support meaning **Second Language Markers** 1. May not be able to demonstrate what he/she knows in a standardized test setting 2. May lack phonemic awareness and concepts of print 3. Very little background knowledge and cultural competence in English, as demonstrated by their reading skills 4. Requires extensive scaffolding and multiple opportunities to interact with familiar text	**Vocabulary** 1. Recognizes words 2. Difficulty with function words (prepositions, adverbs etc.) 3. Works with advanced phonics (blends, root words, prefixes, suffixes) **Structures** 1. Difficulty with complex structures 2. Comprehends basic question formats **Comprehension** 1. Reads at the word and sentence levels 2. Uses graphics support to identify meaning 3. Can make simple inferences 4. Beginning to make connections between words 5. Can make predictions and extend text 6. Can identify the main idea **Strategies** 1. Scans text to match words and phrases **Texts** 1. Reads simple narrative texts **Second Language Markers** 1. Has some background knowledge and cultural competence in English

Level 3	Level 4
Vocabulary 1. Knows some basic vocabulary 2. Recognizes words with similar meanings 3. Comprehends high frequency idiomatic expressions **Structures** 1. Understands passive voice 2. Understands words used as different parts of speech **Comprehension** 1. Compares and contrasts ideas 2. Understands main idea and some supporting details 3. Draws inferences 4. Applies multiple levels of comprehension to literal text 5. Understands basic expository text **Strategies** 1. Emerging use of contextual clues 2. Scans accurately for literal details 3. Uses multiple strategies to identify word meaning **Texts** 1. Reads longer texts for content information 2. Attempts expository 3. Does better with familiar stories and narratives 4. Relies on graphics to support meaning **Second Language Markers** 1. Demonstrates some background knowledge and cultural competence, but not in all areas	**Vocabulary** 1. Use multiple strategies to identify word meanings 2. Can discriminate between closely related vocabulary items **Structures** 1. Understands multiple functions of same words 2. Understands quantifiers and qualifiers **Comprehension** 1. Can disregard or eliminate unnecessary detail 2. Comprehends information from across the text 3. Makes inferences base don information from different parts of the text **Strategies** 1. Applies multiple levels of comprehension to draw an inference 2. Understands explicit and inferred sequence **Text** 1. Sustains comprehension in expository and narrative texts 2. Has ability to read context-reduced materials **Second Language Markers** 1. Has background knowledge and cultural competence in English to interpret narrative and simple expository text

Example of Writing Descriptors for Grade 11

(These levels 1-5 were converted to one landscape 8" x 17" sheet for ease of use with teachers)

Grade 11

Level 1	Level 2	Level 3
Focus/Description 1. Main idea is non-existent or not clear 2. Minimum or no supporting details **Structure/Organization** 1. Disconnected words or phrases 2. Some attempt at sentences 3. Meaning is unclear **Vocab/Semantics/Syntax** 1. Extremely limited vocabulary **Mech/Spell/Punctuation** 1. Simple words spelled correctly **Second Lang. Markers** 1. Requires extensive scaffolding in order to progress through the stages of writing 2. Demonstrates lack of control and under-standing of written academic English forms	**Focus/Description** 1. Lacks a clear focus 2. Fragmented detail **Structure/Organization** 1. Incomplete grasp of basic structures interferes with meaning 2. Simple sentences and/or sentence fragments 3. Words may be omitted 4. Limited sequencing of ideas and events **Vocab/Semantics/Syntax** 1. Limited vocabulary 2. Beginning use of modals **Mech/Spell/Punctuation** 1. Use of phonetic and creative spelling with some basic spelling rules used correctly 2. Use of punctuation, however incorrectly at times. **Second Lang. Markers** 1. Requires extensive scaffolding in order to progress through the stages of writing	**Focus/Description** 1. Generally clear main idea 2. Some supporting details 3. May have weak ending 4. May be prompt or structure dependent 5. Most thoughts are clear and complete **Structure/Organization** 1. Sentence level writing 2. Lacks sense of organization 3. May have fragments and/or run-ons 4. Variety of sentence structures attempted 5. Omission of topic sentences **Vocab/Semantics//Syntax** 1. Informal/familiar register 2. Limited word use 3. Improper vocabulary choices **Mech/Spelling/Punctuation** 1. Emerging use of punctuation and mechanics 2. May use invented and/or phonetic spelling 3. Incomplete control of basic syntax 4. Uses run-ons 5. Demonstrates some under-standing of mechanics and spelling rules **Second Language Markers** 1. Second language markers sometimes interfere with comprehension

Grade 11

Level 4	Level 5
Focus/Description 1. Clear focus and main idea 2. Often uses specific ideas to support main idea **Structure/Organization** 1. Writes in paragraphs 2. Uses multiple verb tenses, but no t always correctly 3. Some variety of sentence sand structures 4. Effective use of transition 5. Clear sense of organization with a beginning, middle and an end 6. Contains few sentence fragments 7. Sentences are logically connected **Vocabulary/Semantics/Syntax** 1. Uses appropriate nouns/verbs and adjectives/adverbs **Mechanics/Spelling/Punctuation** 1. Uses correct grammar most of the time 2. Uses correct punctuation and spelling most of the time **Second Language Markers** 3. Second language errors rarely obscures overall meaning	**Focus/Description** 1. Follows prompt genre 2. Main point defined and maintained 3. Important points/events supported with detail **Structure/Organization** 1. Appropriately paragraphed 2. Uses a variety of sentence lengths and structures 3. Uses a variety of effective transitions 4. Ties the paper together with appropriate ending **Vocabulary/Semantics/Syntax** 1. Uses precise nouns/verbs and adjectives/adverbs 2. May use idiomatic expressions 3. Uses topic-specific vocabulary **Mechanics/Spelling/Punctuation** 1. Predominant use of accurate grammar and spelling **Second Language Markers** 1. Second language errors don't obscure meaning

Appendix C

Frequencies for Language Proficiency Tests

Table C1. Frequencies for LAS-Oral/Listening Scores by Grade

LAS Oral Level (1 low, 5 high)	1		2		3		4		5		Total
	N	%	N	%	N	%	N	%	N	%	
Grade 5th	2	6%	2	6%	8	23%	13	37%	10	28%	35
8th	1	3%	6	19%	19	61%	5	16%	0	0%	31
11th	6	18%	13	39%	12	36%	2	6%	0	0%	33

Table C2. Frequencies for LAS Reading by Grade

LAS Reading Level	1		2		3		Total
	N	%	N	%	N	%	
Grade 5th	5	14%	10	28%	20	57%	35
8th	5	6%	8	26%	18	58%	31
11th	6	18%	13	39%	14	42%	33

Table C3. Frequencies for LAS Writing by Grade

LAS Writing Level	1		2		3		Total
	N	%	N	%	N	%	
Grade 5th	4	11%	28	80%	3	9%	35
8th	3	10%	28	90%	0	0%	31
11th	0	0%	33	100%	0	0%	33

Frequencies for students' scores on the 2003-2004 TEAE were computed for reading and writing. Minnesota uses these performance levels to determine whether a student is English language proficient. Students scoring at level 4 in reading and level 5 in writing are no longer considered ELL for purposes of state funding.

Table C4. Frequencies for TEAE reading Levels by Grade, and by Total Number

TEAE Reading	1		2		3		4		Total
	N	%	N	%	N	%	N	%	
Grade 5th	3	9%	4	11%	23	66%	5	14%	35
8th	2	6%	7	23%	20	65%	2	6%	31
11th	2	6%	16	52%	11	36%	2	6%	31 (2 missing)
Total and % of Total (based on 97 with data)	7	7%	27	28%	54	56%	9	9%	

There were very few students (N=21) in the database with a reported writing score, due in part, to availability of data for urban districts only (Table C5). Percentages overall need to be interpreted with caution because of very small numbers.

Table C5. Frequencies for TEAE Writing Levels by Grade, and by Total Number*

TEAE Writing (urban only)	1		2		3		4		5		Total
	N	%	N	%	N	%	N	%	N	%	
Grade 5th	0	0%	0	0%	0	0%	1	33%	2	67%	3
8th	0	0%	1	20%	1	20%	3	60%	0	0%	5
11th	0	0%	2	16%	4	31%	6	46%	1	7%	13
Total and % of total with data (21)	0	0%	3	14%	5	24%	10	48%	3	14%	

Appendix D

Classroom Rating Data

Table D1. ESL and Content Area Teacher Ratings of Speaking Skills

Speaking	5th N=5		8th N=6		11th N=8	
	ESL	Content	ESL	Content	ESL	Content
5 (Excellent)	1	1	1	-	4	1
4	2	1	2	1	4	5
3	2	3	3	3	-	-
2	-	-	-	2	-	2
1 (Unsatisfactory)	-	-	-	-	-	-

Dash indicates no rating for that number on the scale.

Table D2. ESL and Content Area Teacher Ratings of Listening Skills

Listening	5th N=5		8th N=6		11th N=8	
	ESL	Content	ESL	Content	ESL	Content
5 (Excellent)	0	1	2	2	6	3
4	4	1	3	0	2	3
3	1	2	1	3	0	0
2	0	1	0	1	0	2
1 (Unsatisfactory)	0	0	0	0	0	0

Dash indicates no rating for that number on the scale.

Table D3. ESL and Content Area Teacher Ratings of Writing Skills

Writing	5th N=4		8th N=6		11th N=8	
	ESL	Content	ESL	Content	ESL	Content
5 (Excellent)	-	1	1	1	2	1
4	-	-	3	2	3	2
3	3	1	1	2	3	2
2	1	1	1	1	-	3
1 (Unsatisfactory)	-	1	-	-	-	-

Dash indicates no rating for that number on the scale.

Table D4. ESL and Content Teacher Ratings of Reading Skills

Reading	5th N=5		8th N=6		11th N=7	
	ESL	Content	ESL	Content	ESL	Content
5 (Excellent)	-	-	2	-	2	-
4	2	2	3	1	3	-
3	3	1	1	3	2	3
2	-	1	-	2	-	4
1 (Unsatisfactory)	-	1	-	-	-	-

Dash indicates no rating for that number on the scale.

Table D5. ESL and Content Teacher Ratings of Student Abilities

Student	Grade	Speaking		Listening		Reading		Writing	
		ESL	CON	ESL	CON	ESL	CON	ESL	CON
1	8	3	2	4	3	4	3	4	3
2	8	4	3	5	3	5	2	5	3
3	8	3	3	5	5	4	3	4	4
4	5	3	3	3	2	3	1	2	1
5	5	---	3	---	5	---	4	---	3
6	5	3	5	4	5	4	4	3	5
7	11	4	2	4	2	4	2	4	4
8	11	5	5	5	5	5	3	5	4
9	11	4	4	5	4	3	2	4	2
10	11	5	4	5	4	4	3	4	5
11	11	4	2	4	2	4	2	3	2
12	11	5	4	5	5	5	3	5	3
13	11	4	4	5	5	4	---	3	3
14	11	5	4	5	4	3	2	3-4	2
15	5	3	---	3	---	2	---	2	---
16	5	---	5	---	3	---	3	---	3
17	5	4	4	4	4	3	3	3	3
18	5	---	3	---	4	---	4	---	3
19	5	3	---	2	---	1	---	1	---
20	5	---	2	---	2	---	2	---	2
21	5	4	3	4	3	3	2	3	2
22	8	3	2	3	2	3	2	2	2
23	8	4	3	4	5	5	4	3	4
24	8	5	4	4	3	4	3	4	5
25	5	5	3	4	3	4	4	---	2

Dash indicates no rating.

Table D6. ESL and Content Teacher Ratings of Student Performance in Relation to Their Fluent English Peers

	5th N=4		8th N=6		11th N=8	
	ESL	Content	ESL	Content	ESL	Content
(Excellent) **5**	-	1	-	1	1	2
4	-	-	-	2	-	2
3	1	1	-	3	-	1
2	1	2	-	-	-	1
1 (Unsatisfactory)	-	-	-	-	-	0
NA *	2		6		7	2

Dash indicates no rating for that number or scale.
* Indicates no English fluent peers in class.

Table D7. ESL and Content Teacher Ratings of Student Abilities on Additional Questions

Student	**Grade**	**Ability to Master Content**		**Performance Relative to Peers**		**Chance of Success without ESL assistance**	
		ESL	CON	ESL	CON	ESL	CON
1	8	5	4	NA	3	4	4
2	8	5	3	NA	3	5	2
3	8	5	5	NA	4	4	4
4	5	2	2	2	2	3	2
5	5	---	3	---	2	---	2
6	5	3	5	3	5	3	5
7	11	4	4	NA	NA	2	2
8	11	5	5	NA	5	4	5
9	11	5	5	5	4	3	5
10	11	5	5	NA	4	4	3
11	11	3	3	NA	3	3	2
12	11	5	3	NA	2	4	2
13	11	4	5	NA	5	3	5
14	11	5	5	NA	NA	1-2	1
15	5	3	---	3	---	2	---
16	5	---	5	---	3	---	3
17	5	3	3	NA	3	2	2
18	5	---	4	---	4	---	5
19	5	1	---	1	---	1	---
20	5	---	2	---	2	---	2
21	5	4	2	NA	2	3	1
22	8	3	2	NA	3	2	2
23	8	5	5	NA	5	5	5
24	8	4	4	NA	4 -5	4	4-5
25	5	5	2	---	2	4	2

Dash indicates no rating.

Table D8. Ratings Provided by ESL Teachers Using State-Developed Descriptors

Student Sample	Teacher		Rater 1		Rater 2		Rater 3	
	R	W	R	W	R	W	R	W
1	3	4	3	3	3	4		3
2	4	4-5	2-3	4	3	4		
3	3-4	4	2-3	3	3	4	3-4	3-4
4	2	2	3	3	2	4		
5	---	---	---	---	---	---		
6	3	3	2-3	3	3	3		
7	3*	3	---	3	---	3		
8	4	5	4	4	4	4		4
9	3	4	---	3	---	2		
10	3*	4	---	3-4	---	4		
11	4	4	3-4	3	3	3	3	3
12	3-4	3	3	3	3	3		
13	3	4	3	3	3	4		3
14	---	3	2-3	2	2	3		
15	2	2	2-3	2	3	2		
16	---	---	4	4	3-4	4		
17	---	3	---	3	---	2		
18	---	---	---	---	---	---		
19	---	1	---	1-2	---	2		
20	---	---	---	---	---	---		
21	3	3	3	3-4	2-3	2-3		
22	2	2	1	2	2	2-3		
23	3	4	3	3	2-3	3		
24	2-3	4	2	3	2	3		
25	3	3	2-3	2-3	4	3		

Note: * Denotes no sample provided. --- Denotes no rating provided by the rater.

Table D9. Ratings Provided by Content Teachers Using State-Developed Descriptors

Student Sample	Teacher		Rater 1		Rater 2		Rater 3	
	R	W	R	W	R	W	R	W
1 (copied)	3	4	3	3	3	4		
2 (copied)	3	4	3	3	3	4		
3	2	3	3	3	2	---	3-4	
4	3	---	2-3	3	3	4		4
5	2-3	1	2-3	2	3	2		3
6	4	5	3	4	4	5	3	
7	2	3-4	2	3	3	3		
8	4	5	3-4	3-4	3	4	---	4
9	3	3	---	3	2	3		
10	3	4	3	3	3	4		
11	4	4	3	3	3	3	3	4
12	2	3	3	3	3	3	3	
13	4	3	3	3	3	3		
14	3	2	2-3	2-3	3	2		
15	---							
16	---							
17	1-2	2	3	2	3	3		
18	4	3	3	3	3	3-4	3	
19	---							
20	2	3	3	3-4	2-3	3-4		
21	4	3	3	4	2-3	3	2	
22	---	4	---	3	---	3		
23	---	---	---	3-4	---	3		
24	---	3-4	---	3	---	3		
25	3-4	3	2-3	1-2	3-4	3-4		

Note: * Denotes no sample provided. --- Denotes no rating provided by the rater.

Table D10 was created using the individual student rating data in Tables D7 and D8.

Table D10. Percentage of Agreement between Raters on Work Sample Rubric Rating

Work Sample Provided by		Type of Comparison					
		Teacher to Rater 1	Teacher to Rater 2	Rater 1 to Rater 2	Teacher to Rater 3	Rater 1 to Rater 3	Rater 2 to Rater 3
	ESL Teacher	68%	68%	78%	0%	100%	71%
	Content Teachers	54%	63%	69%	11%	56%	56%

Table D11. ESL Teacher Reading Ratings, Averages, and Standard Deviations on Standardized Tests

	ESL Teacher Rating of Student Reading Ability			
	1 (N=0)	**2** (N= 4)	**3** (N= 10)	**4** (N = 3)
TEAE Reading Scaled Score (and SD) N	 0	214.25 (18.30) N = 4	235.80 (32.40) N = 10	276.00(46.77) N = 3
MCA Reading Score (and SD) N	 0	1425.00 (49.50) N = 2	1503.33 (243.38) N = 3	 N = 0
BST Reading Score (and SD) N	 0	542.00 (33.94) N = 2	605.80(14.02) N = 5	620.33(32.13) N = 3

Table D12. Content Teacher Reading Ratings, Averages, and Standard Deviations on Standardized Tests

	Content Teacher Ratings of Students' Reading Ability			
	1 (Total[1] N=1)	**2** (Total N= 5)	**3** (Total N= 8)	**4** (Total N = 5)
TEAE Reading Scaled Score (SD) N	222.00 1	225.60 (26.15) 5	233.38 (37.15) 8	246.60 (47.82) 5
MCA Reading Score (SD) N	1410.00 1	1280.00 1	1470.00 (113.14) 2	1476.67 (240.07) 3
BST Reading Score (SD) N	 0	617.00 (0.00) 2	594.50 (23.81) 4	622.00 (45.25) 2

[1]Total N is for all students that had rating data.

Table D13. ESL Teacher Writing Ratings, Averages, and Standard Deviations on Standardized Tests

	ESL Teacher Rating of Writing Ability				
	1 (N=1)	**2** (Total[1] N = 3)	**3** (Total N = 7)	**4** (Total N =8)	**5** (Total N = 1)
TEAE Writing Score (and SD) N	 0	17.0 1	20.67 (2.08) 3	26.00 (0.00) 3	29.00 1
MCA Writing* Score (and SD) N	 0	1485.00 (176.78) 2	1607.50 (192.59) 4	 0	 0

[1]Total N is for all students that had rating data.
*There were no published writing levels available from the state for interpreting mean writing scores shown for writing.

Table D14. Content Teacher Writing Ratings, Averages, and Standard Deviations on Standardized Tests

	Content Teacher Ratings of Students' Writing Ability				
	1 (Total[1] N=1)	**2** (Total N= 2)	**3** (Total N= 9)	**4** (Total N=6)	**5** (Total N = 2)
TEAE Writing Score (SD) N	0	19.00 1	22.60 (3.58) 5	21.50 (6.36) 2	29.00 1
MCA Writing* Score (SD) N	0	1470.00 1	1642.50 (245.68) 4	0	1840.00 1

[1]Total N is for all students that had rating data.
*There were no published writing levels available from the state for interpreting mean writing scores shown for writing.